CANCER IN YOUNG PEOPLE IN CANADA PROGRAM

PROGRAMME CANCER CHEZ LES JEUNES AU CANADA

GIS for CYP-C Research

Based on research by the lab of Dr. Alvaro Osornio-Vargas

Presented by Charlene Nielsen [ccn@ualberta.ca]

UNIVERSITY OF ALBERTA
FACULTY OF MEDICINE & DENTISTRY
Department of Pediatrics

Outline

- Overview of GIS
- Basemaps and map projections
- Geolocation
- Analyses: statistics, overlay, and proximity
- Other learning references

What a GIS is **not**

GPS

Global Positioning System

GPS data can be used in GIS analyses

Static Map

Digital/paper map is an "input" or "product" of GIS

A way to visualize output from GIS analyses

Software

Functions and tools needed to store, analyze, and display geographic information

Requires hardware, data, and personnel in a complete system

Database

Set of tables containing data that can be accessed or reassembled in many different ways

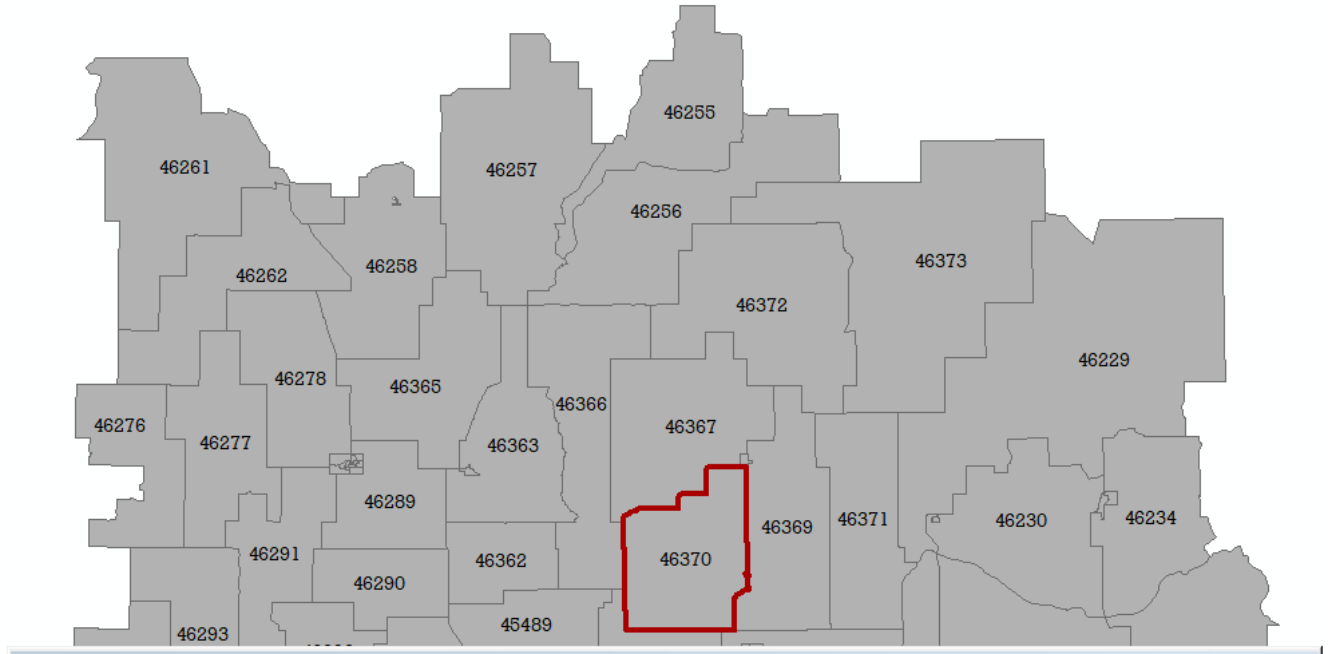
Requires a link to spatial data

UNIVERSITY OF ALBERTA
FACULTY OF MEDICINE & DENTISTRY
Department of Pediatrics

Geographical Information Systems

UNIVERSITY OF ALBERTA
FACULTY OF MEDICINE & DENTISTRY
Department of Pediatrics

Link databases and maps



Table

Dissemination Areas

OBJECTID	Shape *	DAUID	CDUID	TotalPop	ChildrenPop	SES	Type	AREASQKM	ChildDensity
46367	Polygon	48130309	4813	478	135	0.55	2	288.73	0.47
46368	Polygon	48130310	4813	505	85	0.07	2	1.68	50.74
46369	Polygon	48130311	4813	486	115	-0.0	2	244.2	0.47
46370	Polygon	48130312	4813	498	100	0.17	2	229.75	0.44
46371	Polygon	48130313	4813	528	130	0.04	2	223.77	0.58
46372	Polygon	48130314	4813	504	130	-0.1	2	384.36	0.34

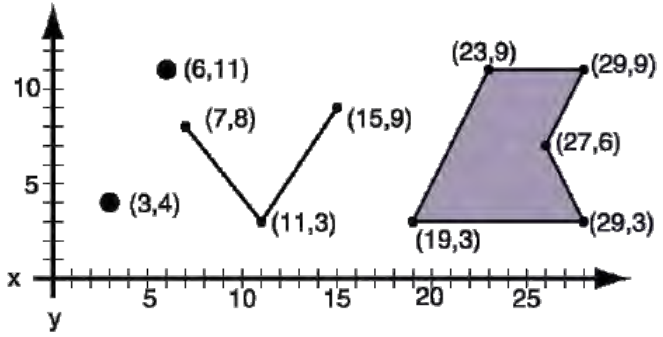
0 (1 out of 140 Selected)

Dissemination Areas

Types of spatial data

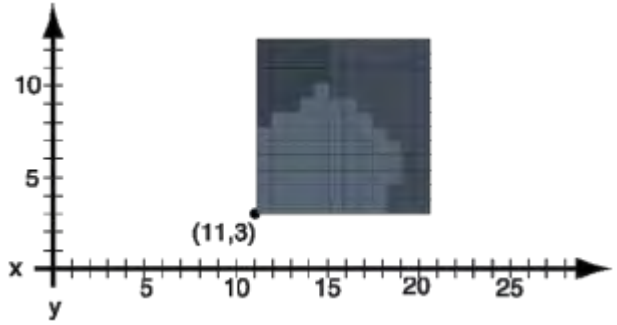
Vector

Models discrete features as points (coordinates), lines (arcs), and polygons (areas) with precise boundaries and shapes with attributes; e.g. feature class, shapefile, CAD drawing

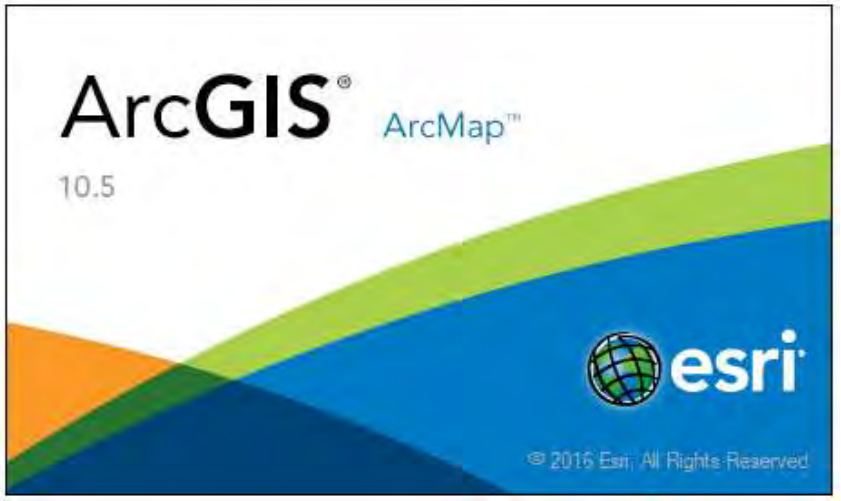


Raster

Models continuous phenomena in a surface divided into a regular grid of cells (pixels) each having an associated attribute value; e.g. grid coverage, TIFF, digital photo



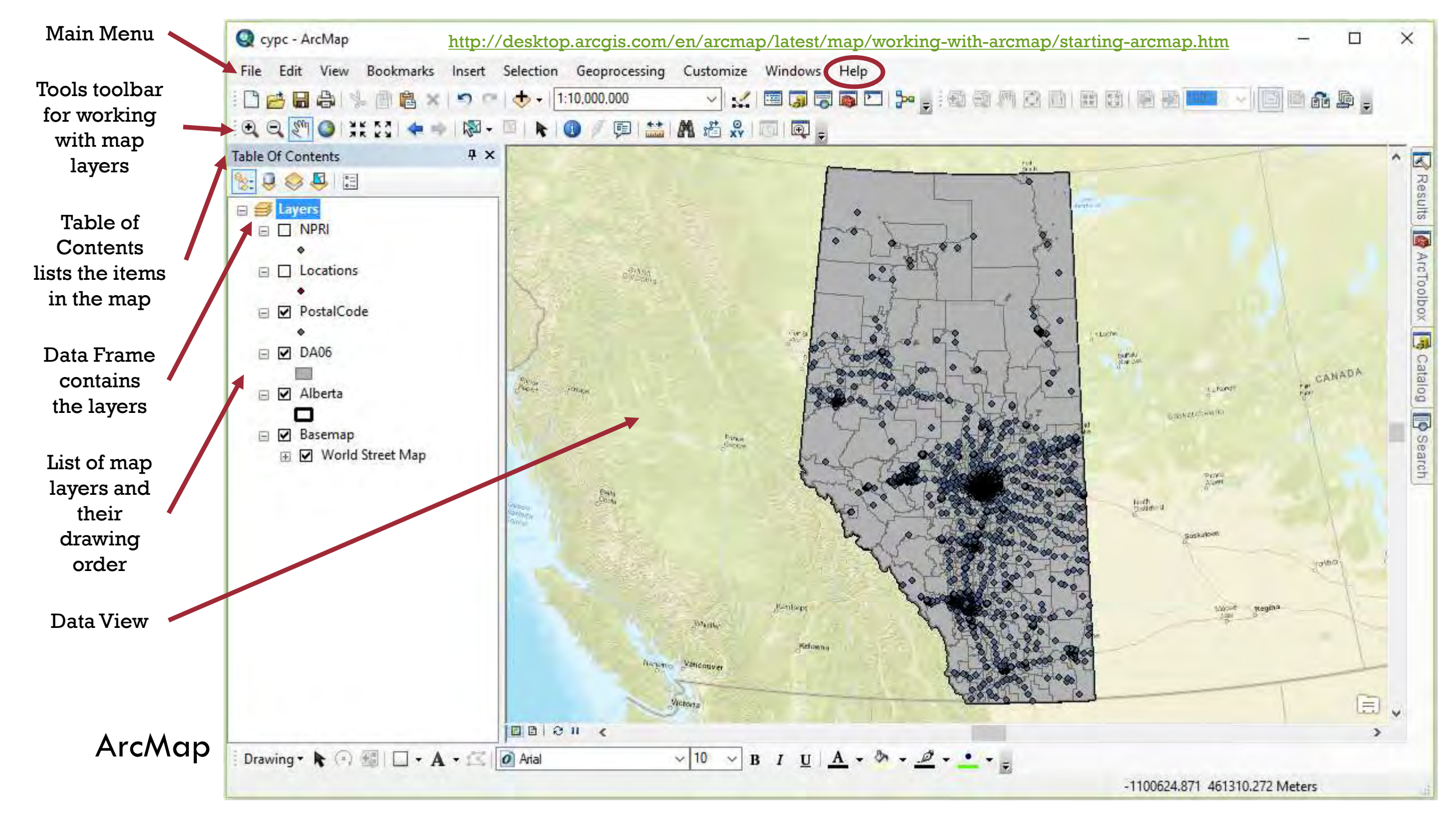
ArcGIS Desktop

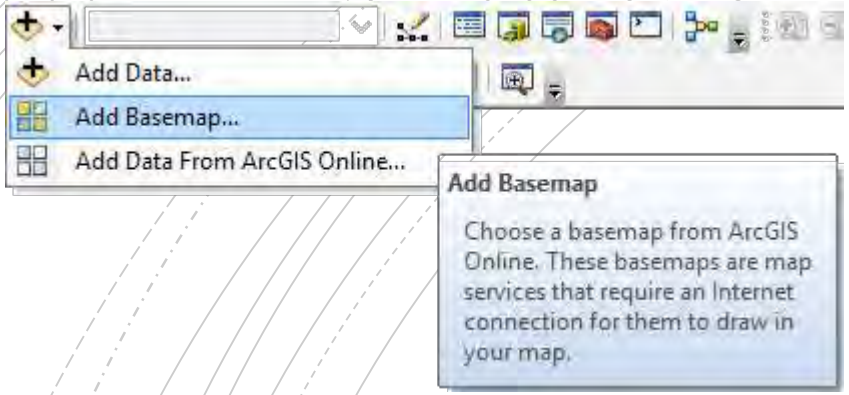


http://www.esri.com/en/arcgis/products/arcgis-pro/Resources/ArcMap-Resources

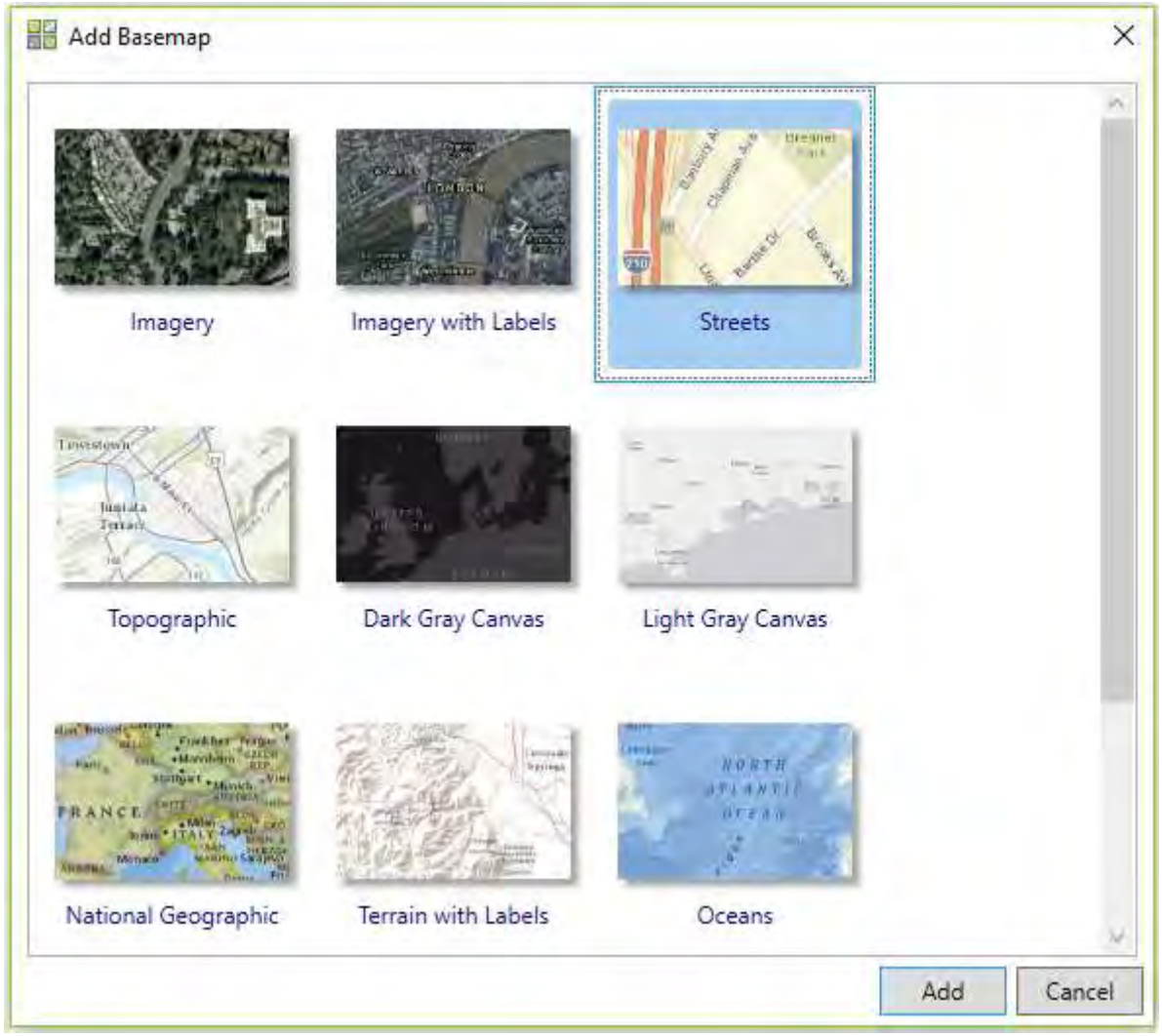
"Powerful desktop applications create maps, perform spatial analysis, and manage data in 2D and 3D."

"ArcMap is the industry leading, traditional GIS authoring and editing application."


Main Menu
Tools toolbar for working with map layers
Table of Contents lists the items in the map
Data Frame contains the layers
List of map layers and their drawing order
Data View
ArcMap
cypc - ArcMap
http://desktop.arcgis.com/en/arcmap/latest/map/working-with-arcmap/starting-arcmap.htm
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
1:10,000,000
Table Of Contents
Layers
NPRI
Locations
PostalCode
DA06
Alberta
Basemap
World Street Map
CANADA
Results
ArcToolbox
Catalog
Search
Drawing
Arial
10
-1100624.871 461310.272 Meters



Basemaps



http://links.esri.com/arcgisdesktop/en/arcmap/layers

Navigate

Zoom In

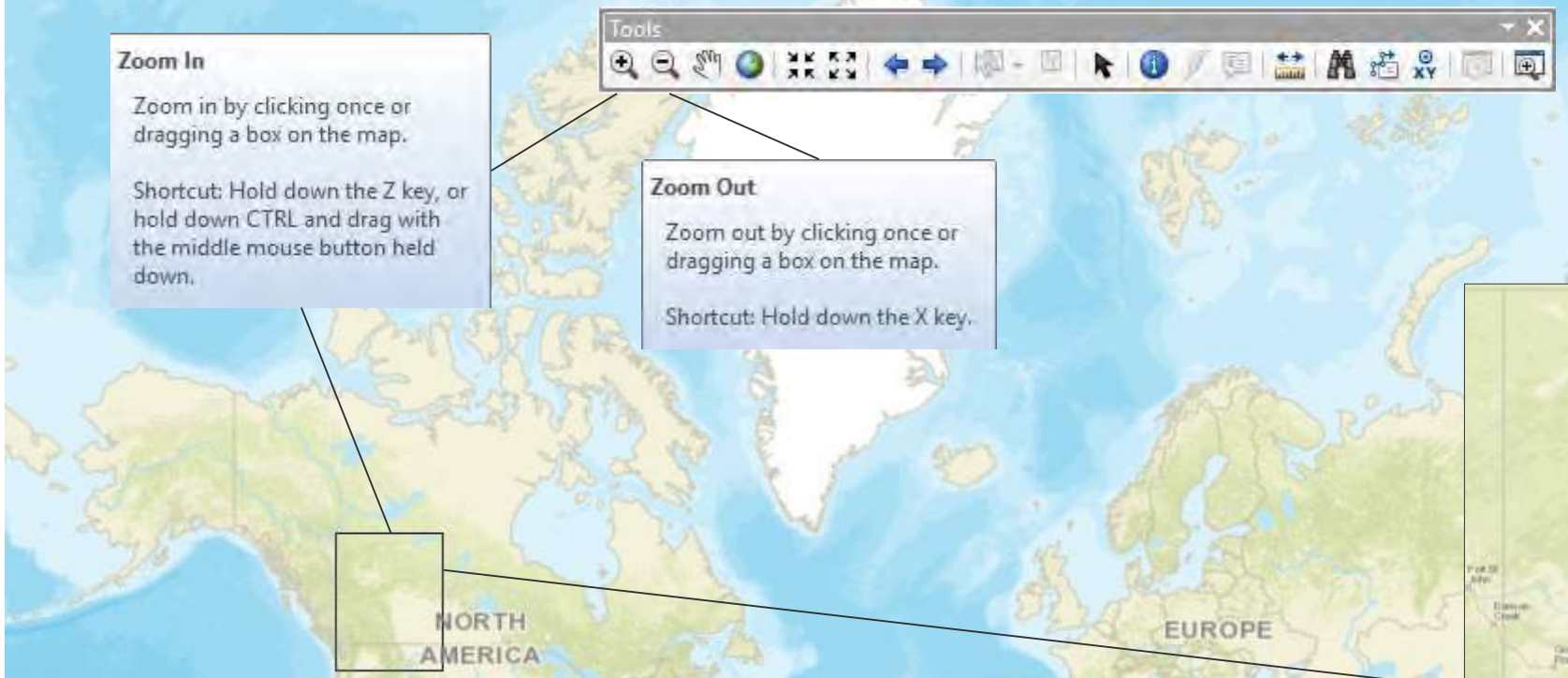
Zoom in by clicking once or dragging a box on the map.

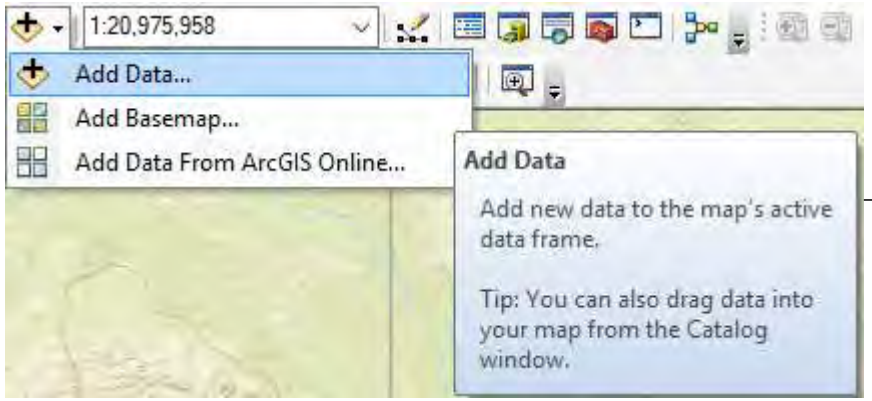
Shortcut: Hold down the Z key, or hold down CTRL and drag with the middle mouse button held down.

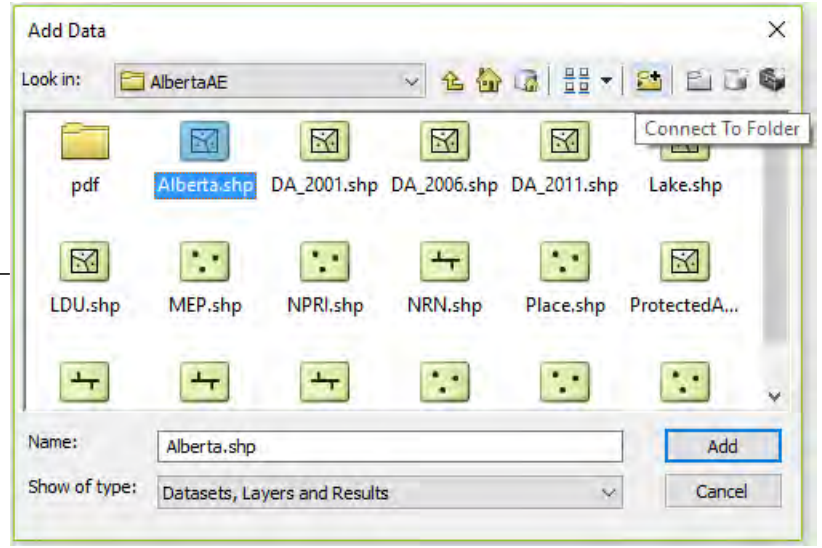
Zoom Out

Zoom out by clicking once or dragging a box on the map.

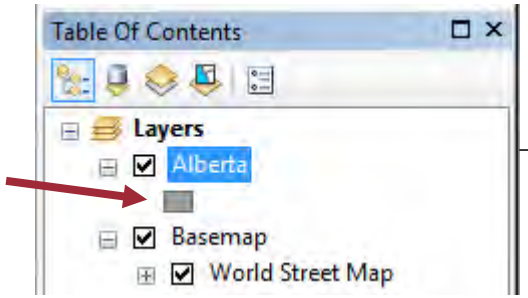
Shortcut: Hold down the X key.

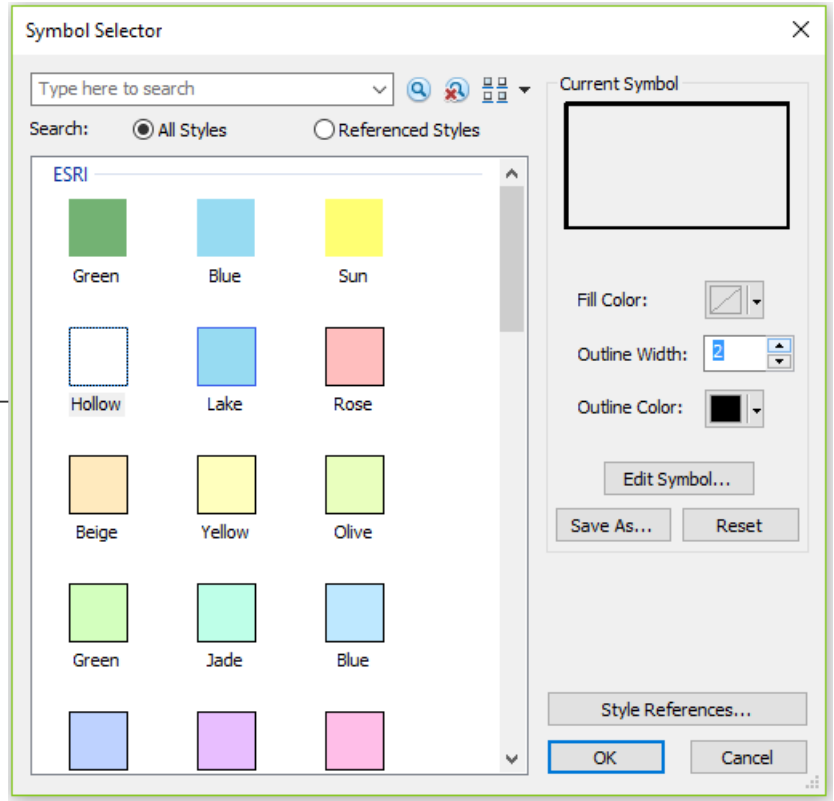




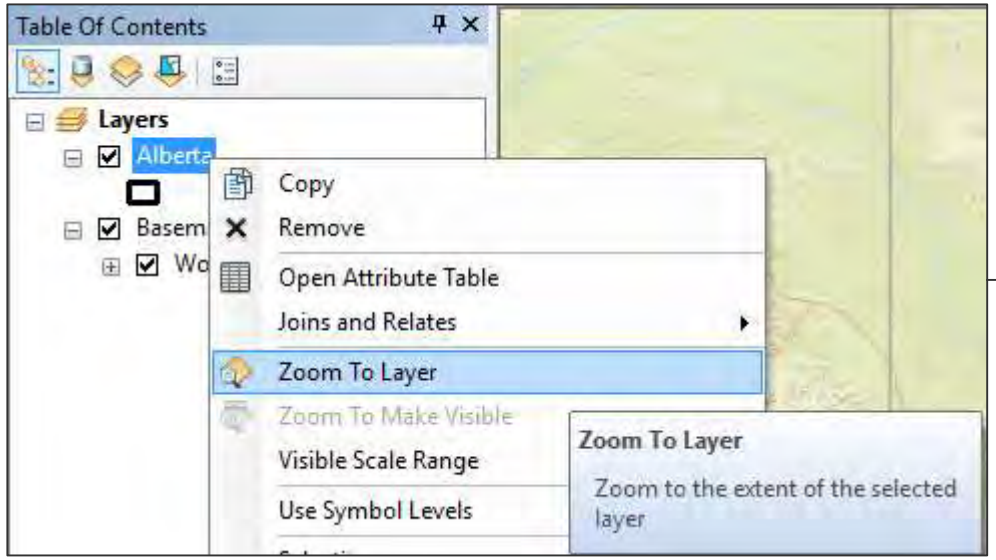


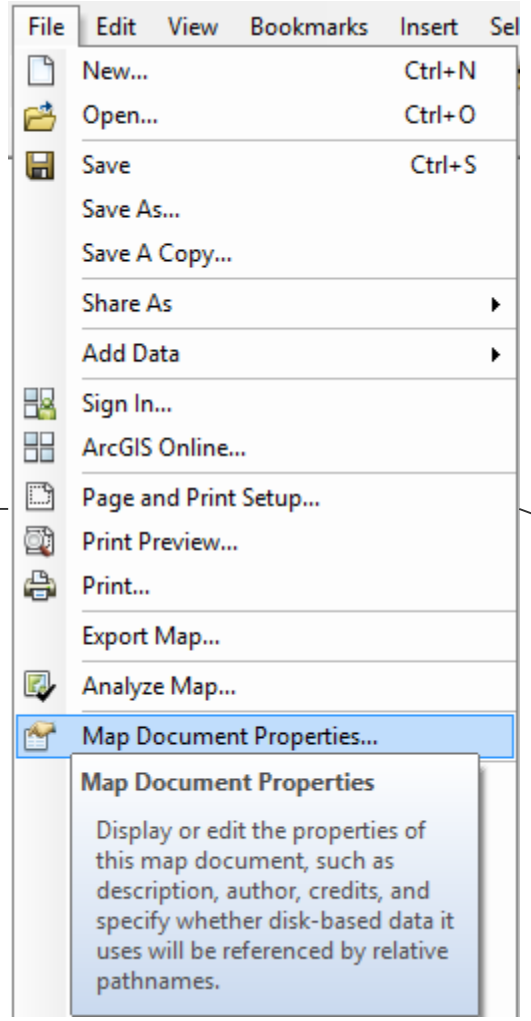
Symbol patch

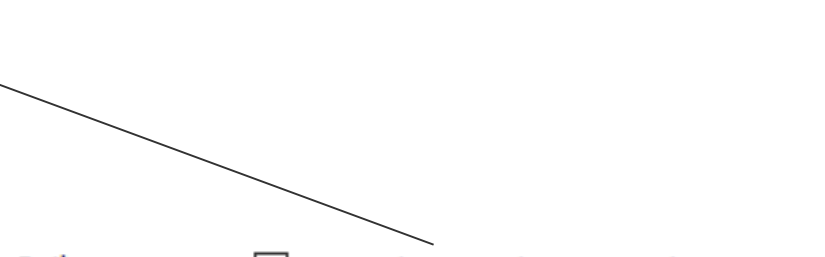


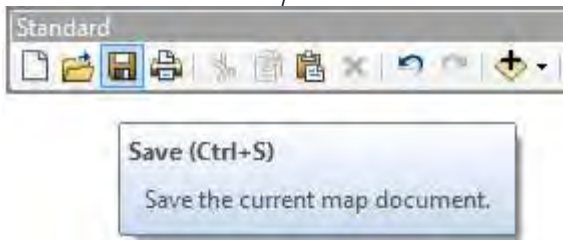


Add data and symbolize







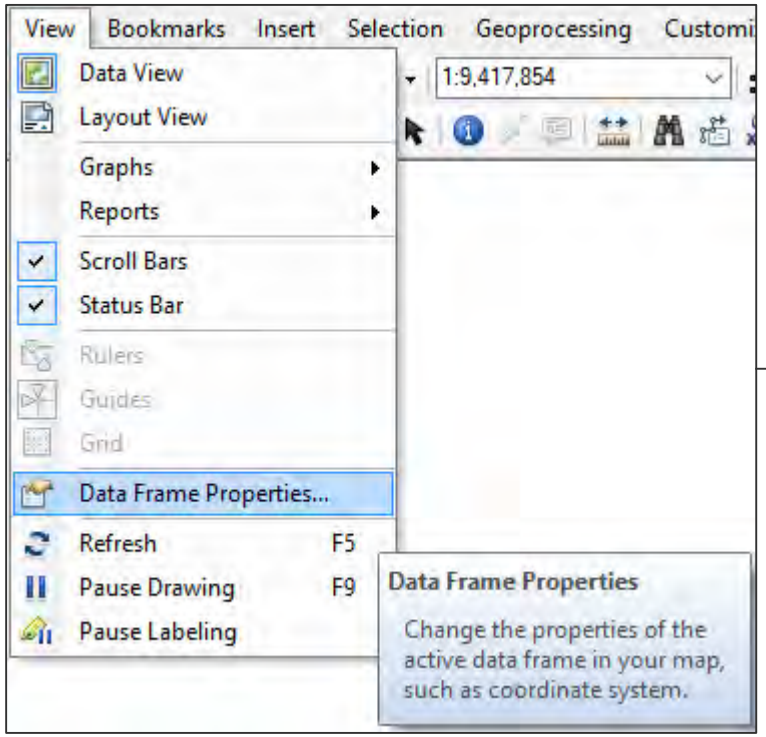


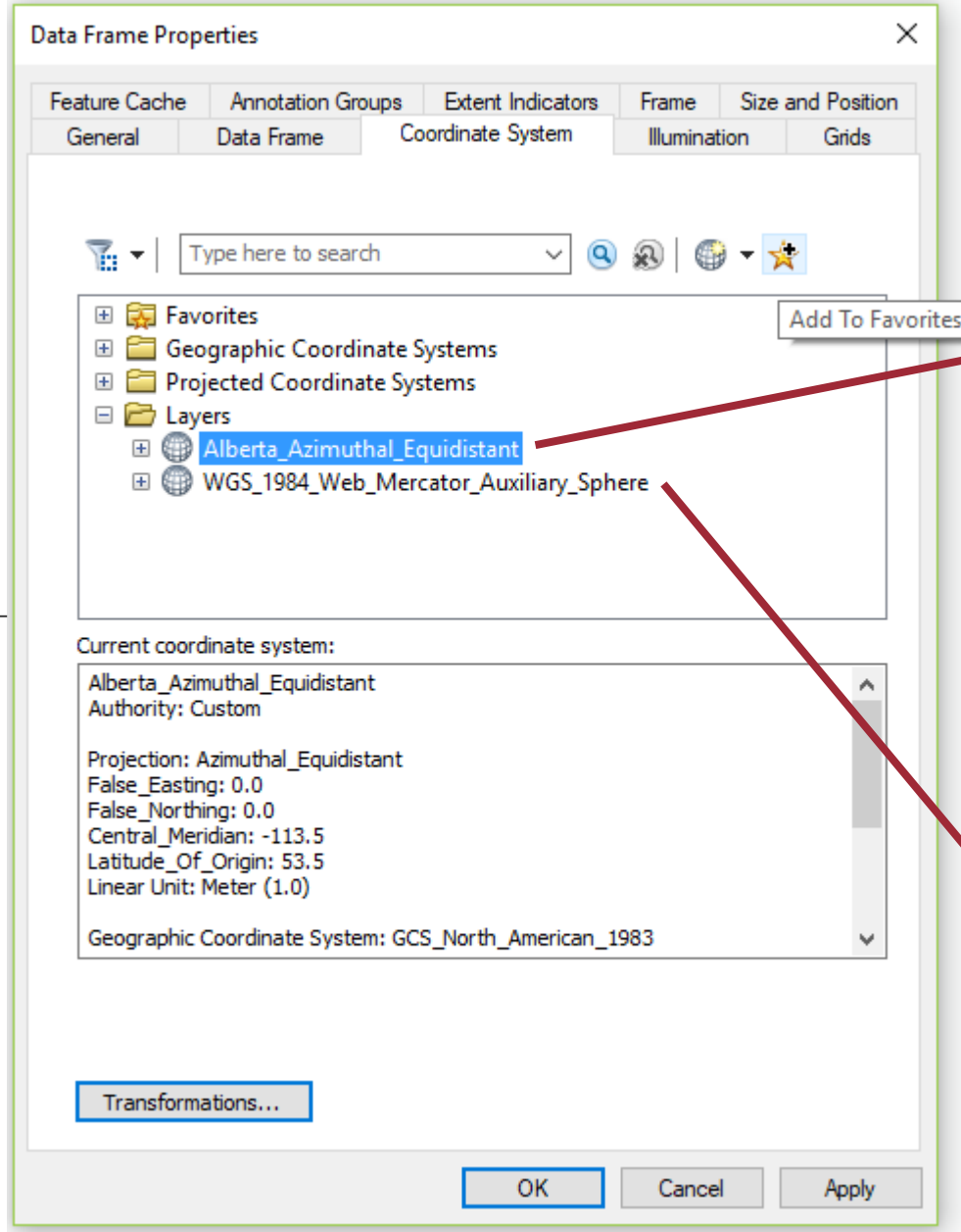
Zoom and save

Map projections

Name	False Easting	False Northing	Central Meridian	Standard Parallel 1	Standard Parallel 2	Latitude of Origin
NAD 1983 Canada Atlas Lambert	0	0	-95	49	77	49
Alberta Azimuthal Equidistant	0	0	-113.5	n/a	n/a	53.5
NAD 1983 10TM AEP Forest	500000	0	-115	n/a	n/a	0
NAD 1983 BC Environment Albers	1000000	0	-126	50	58.5	45
NAD 1983 Ontario MNR Lambert	930000	6430000	-85	44.5	53.5	0
NAD 1983 Quebec Albers	0	0	-68.5	46	60	44

http://desktop.arcgis.com/en/arcmap/latest/map/projections/what-are-map-projections.htm





Coordinate system

Geolocation

GIS Dictionary

Look up terms related to GIS operations, cartography, and Esri technology.

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Back to list

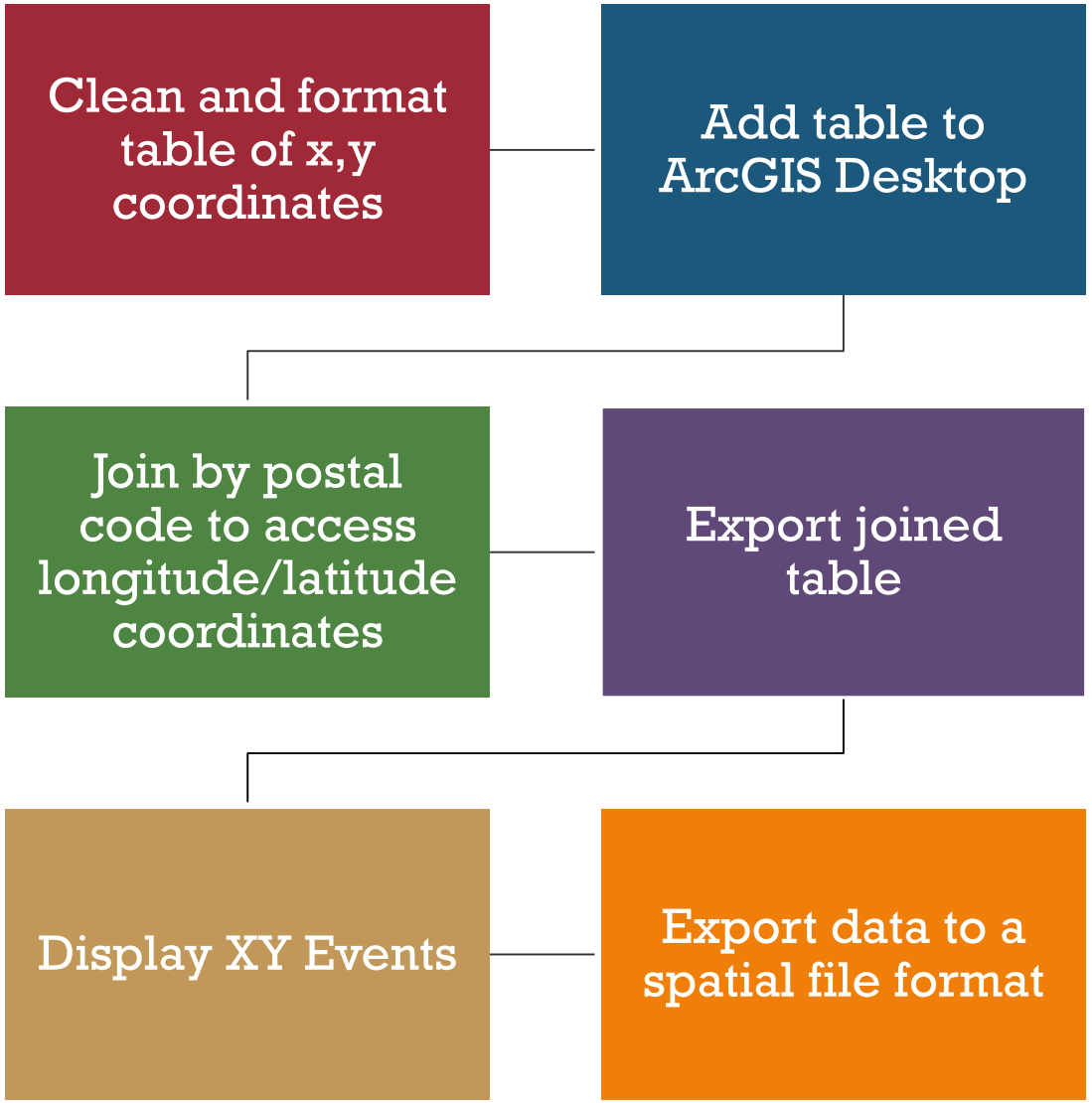
http://support.esri.com/en/other-resources/gis-dictionary

geolocation

1 [geolocating] The process of creating geographic features from tabular data by matching the tabular data to a spatial location. An example of geolocation is creating point features from a table of x,y coordinates.

UNIVERSITY OF ALBERTA
FACULTY OF MEDICINE & DENTISTRY
Department of Pediatrics

Geolocation



Standardize the Attributes

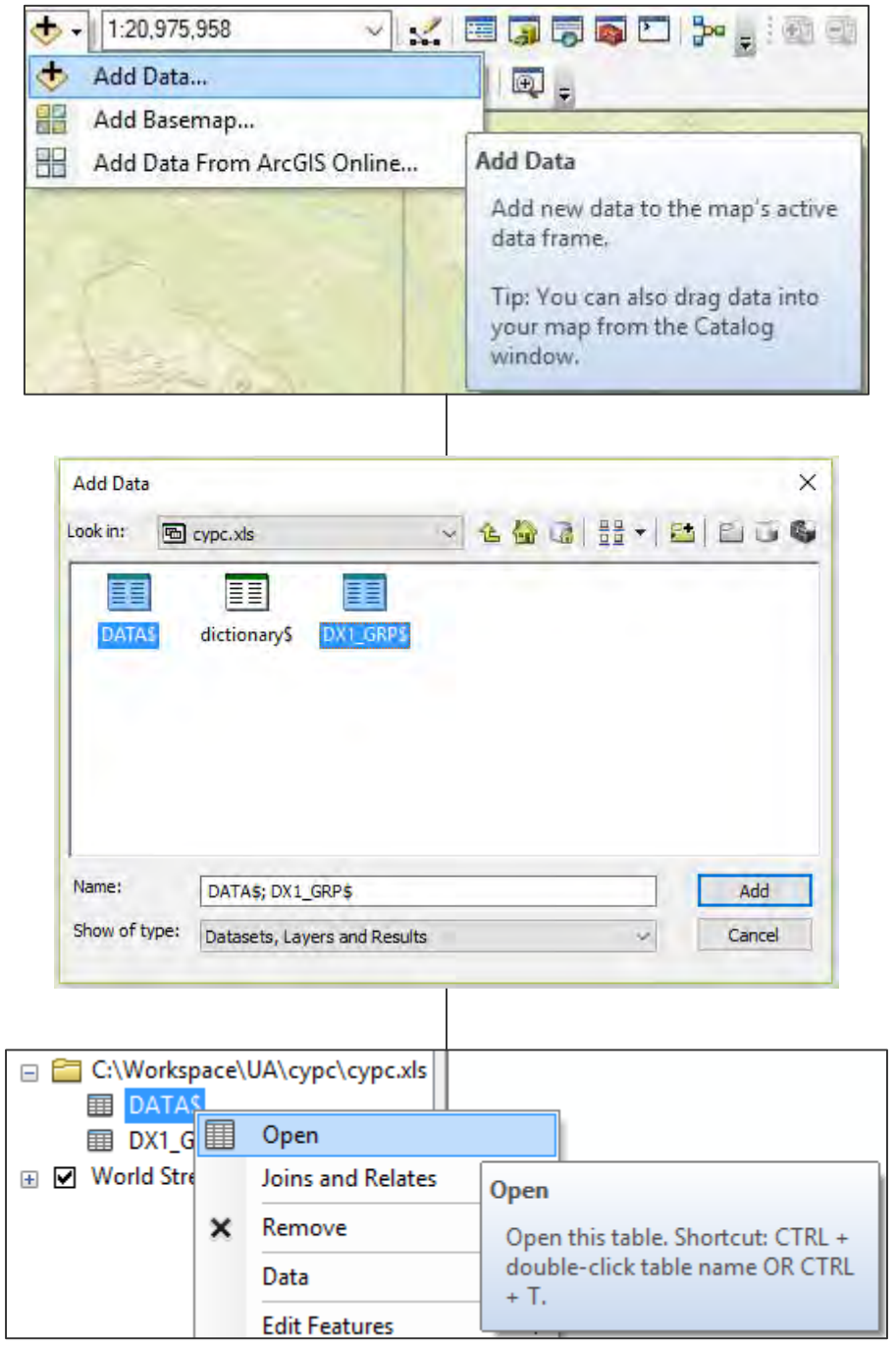
- No database is initially ever "clean" and getting your data ready for spatial analysis will take upwards of 80% of your time
- Full 6-character postal code
  - CAPitaLIZaTioN matters
    T6G 1c9 ≠ T6G 1C9
  - Ensure all are in the format of A1A 1A1
    T0L OC0 ≠ T0L 0C0
  - Remove all spaces
    T6G1C9_ ≠ T6G1C9

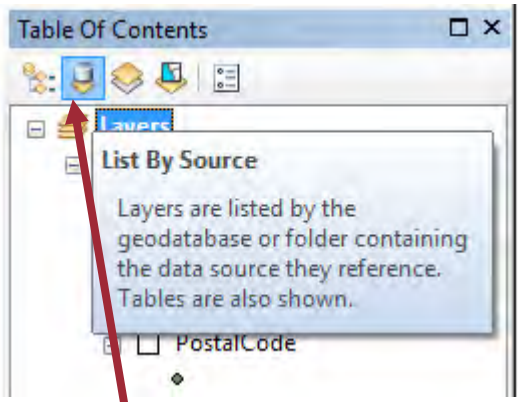
Tips for Tables

- Column headings must be present
  - No non-alphanumeric characters (e.g. %)
  - Do not use spaces – check at end of the word
- No skipped rows anywhere
- Date/time values are subject to import errors
  - May split the date parts in to separate columns
- Export table from your statistical software to: *.dbf, *.csv, or *.xls
  - ArcGIS can only read MS Excel version 2003 or earlier (no *.xlsx files)

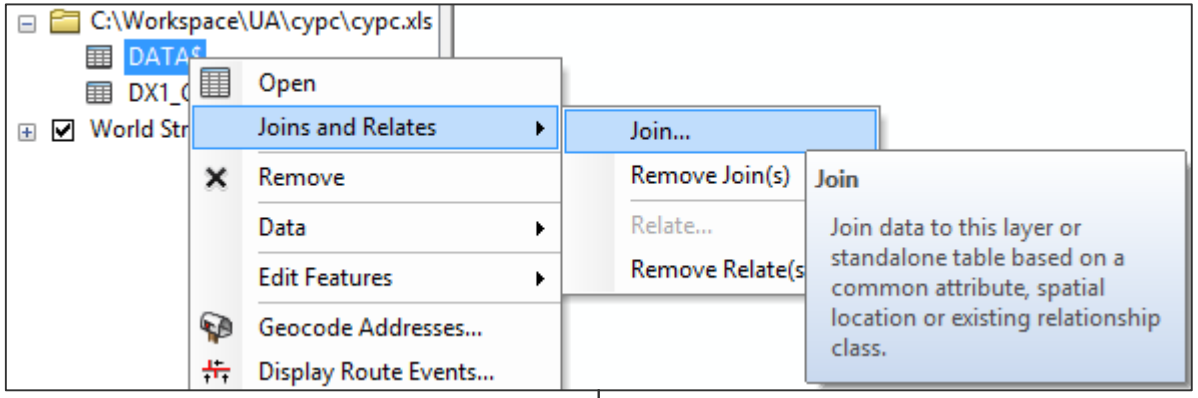
Clean and format

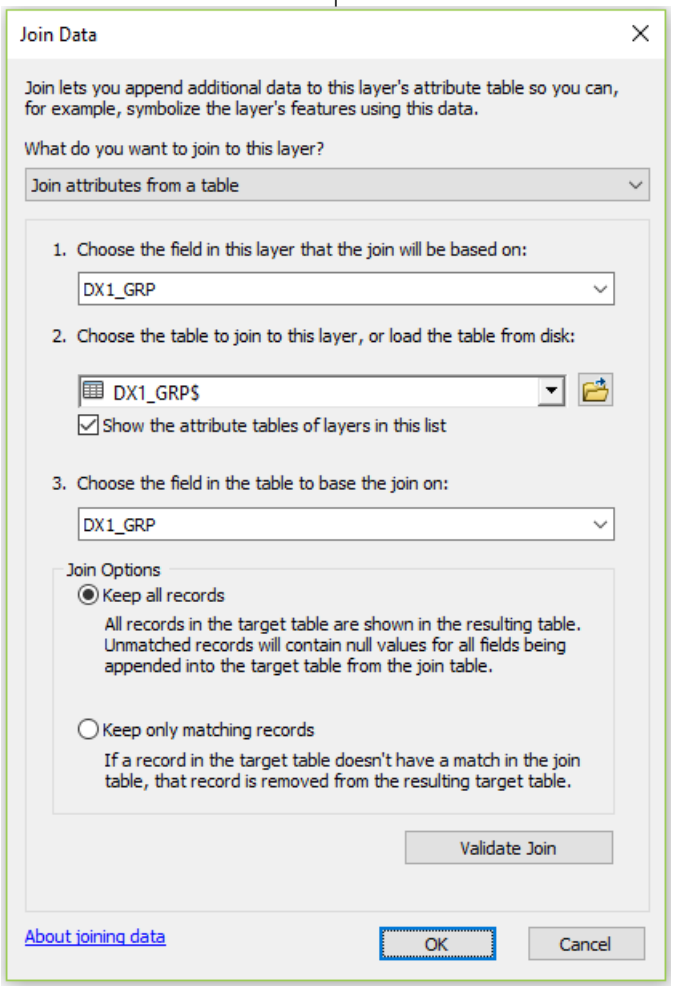
Add data

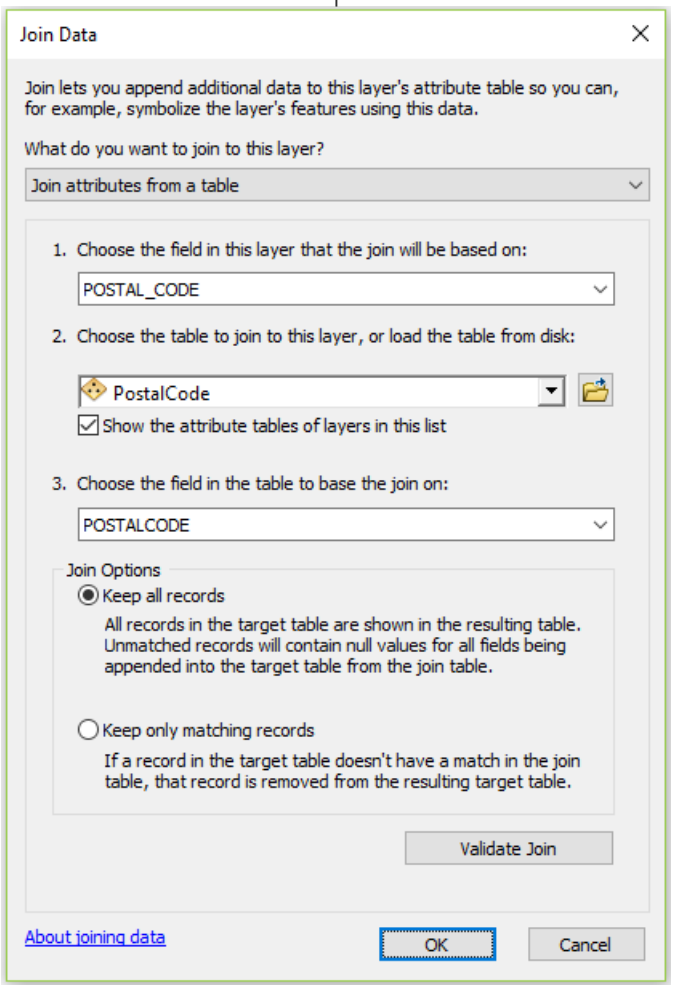

1:20,975,958
Add Data...
Add Basemap...
Add Data From ArcGIS Online...
Add Data
Add new data to the map's active data frame.
Tip: You can also drag data into your map from the Catalog window.
Add Data
Look in:
cypc.xls
DATA$
dictionary$
DX1_GRP$
Name:
DATA$; DX1_GRP$
Add
Show of type:
Datasets, Layers and Results
Cancel
C:\Workspace\UA\cypc\cypc.xls
DATA$
DX1_G
World Stre
Open
Joins and Relates
Remove
Data
Edit Features
Open
Open this table. Shortcut: CTRL + double-click table name OR CTRL + T.



To see stand-alone tables in the Table of Contents, click the List By Source icon







Join

The DATA$ table joined to DX1_GRP$ (to access the diagnosis description) and PostalCode (to access the coordinates) tables

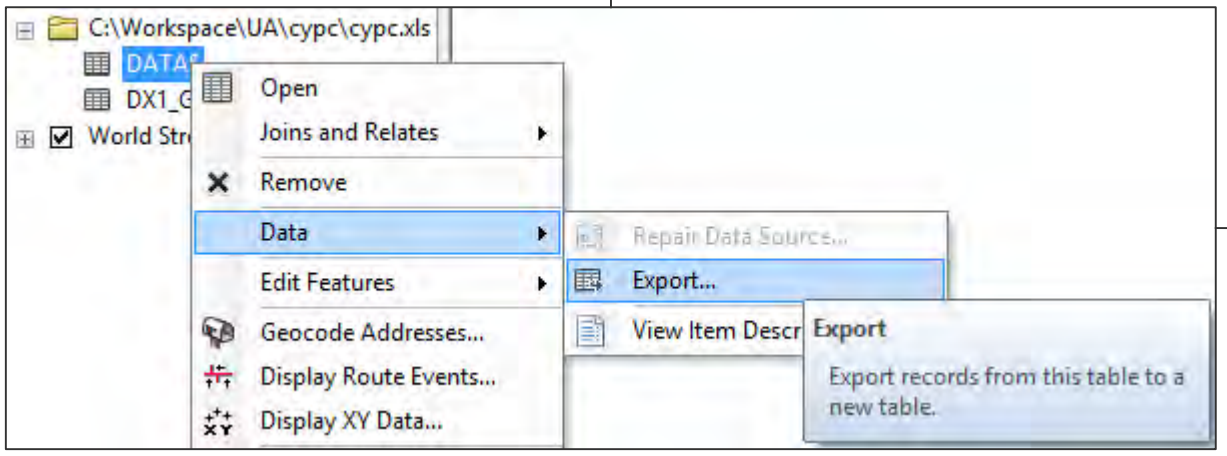
Table

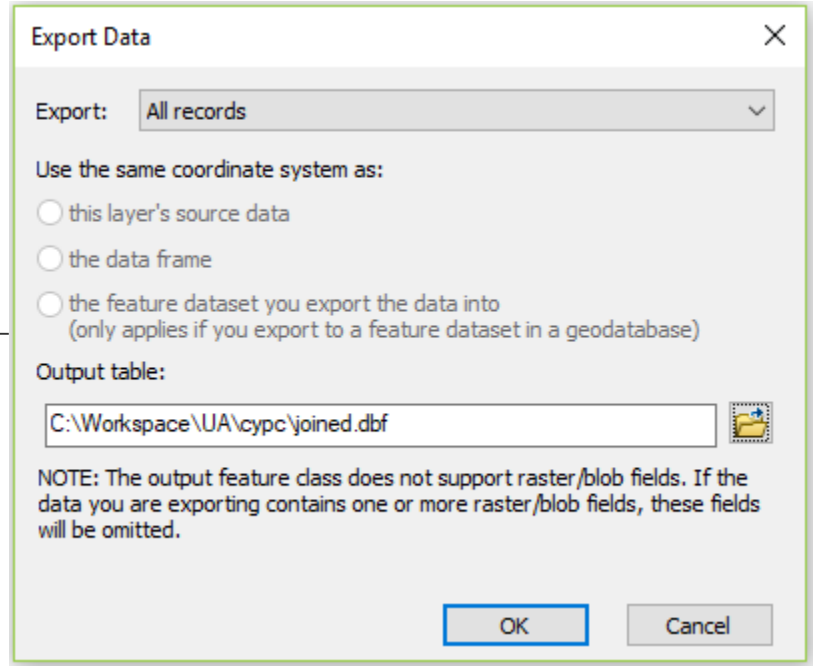
DATA$

CYPCID	DOB	GENDER	POSTAL_CODE	DX_DATE	DX_AGE	DX1_GRP	CHEMO	DX1_GRP	Diagnosis_Group	FID	POSTALCODE	LONGITUDE	LATITUDE	DAUID
1	1999-03-20	M	T1A2L4	2004-03-20	5	11	1	11	Hepatic tumors	5513	T1A2L4	-110.683259	50.029562	48010199
2	1997-04-18	F	T2T3E5	2005-04-18	8	6	0	6	Renal tumors	18534	T2T3E5	-114.084688	51.02591	48060617
3	2004-08-10	M	T2T3P9	2015-08-10	11	9	0	9	Other malignant epithelial neoplasms and malignant melanomas	18191	T2T3P9	-114.092444	51.023647	48060619
4	1999-03-30	F	T1A4K1	2002-03-30	3	5	0	5	Soft tissue and other extraosseous sarcomas	5680	T1A4K1	-110.692985	50.032527	48010194
5	1997-05-31	F	T6B0W7	2008-05-31	11	5	1	5	Soft tissue and other extraosseous sarcomas	63082	T6B0W7	-113.436916	53.529479	48110821
6	1998-08-12	M	T9H2V4	2012-08-12	14	12	0	12	Other and unspecified malignant neoplasms	87789	T9H2V4	-111.364116	56.698394	48160122

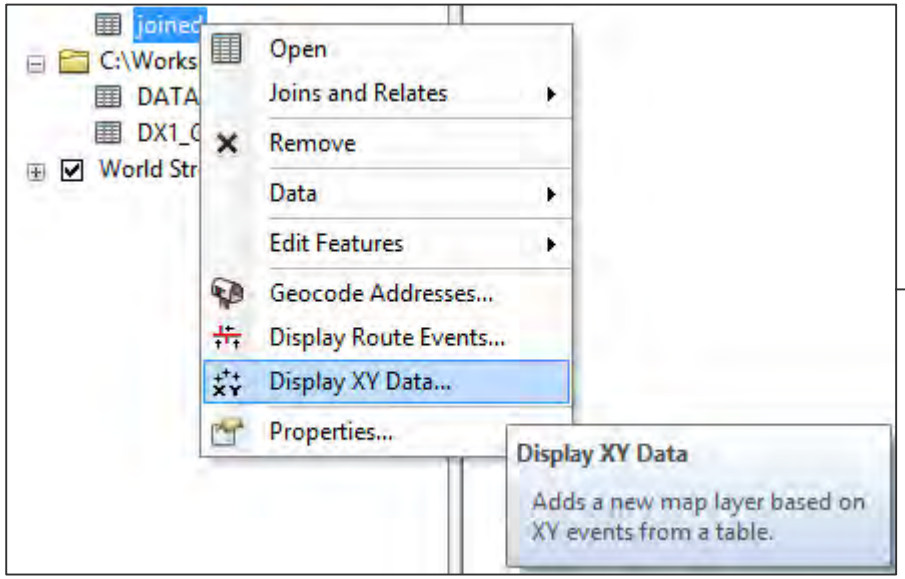
1 (0 out of 333 Selected)

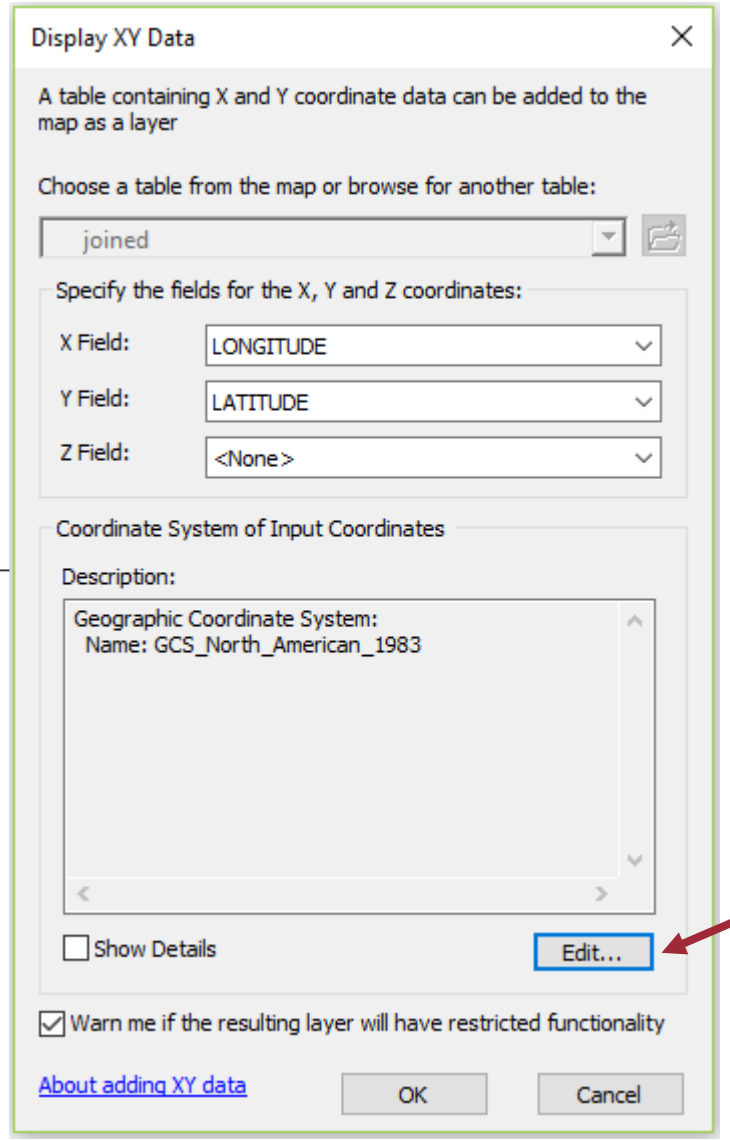
DATA$





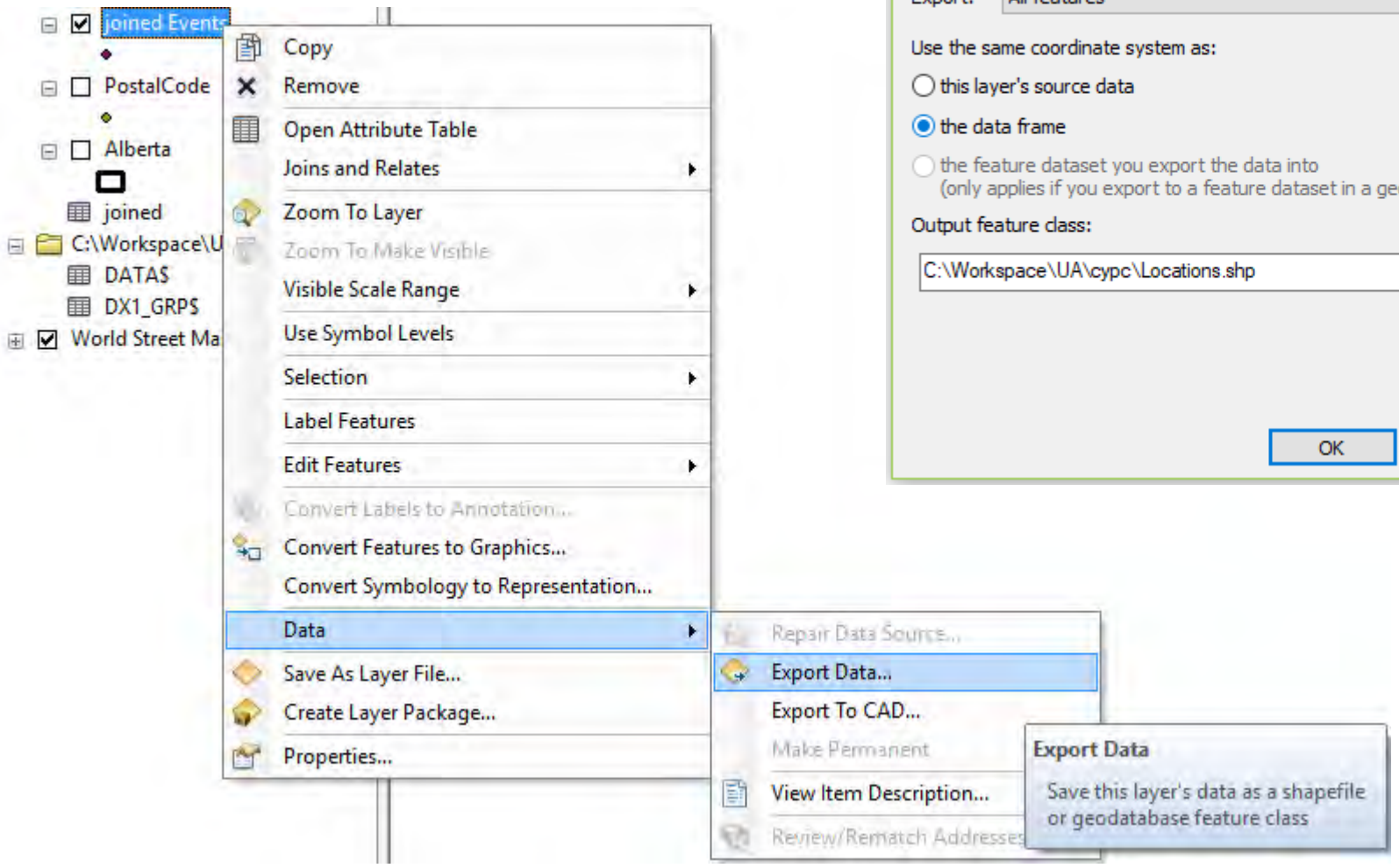
Export table

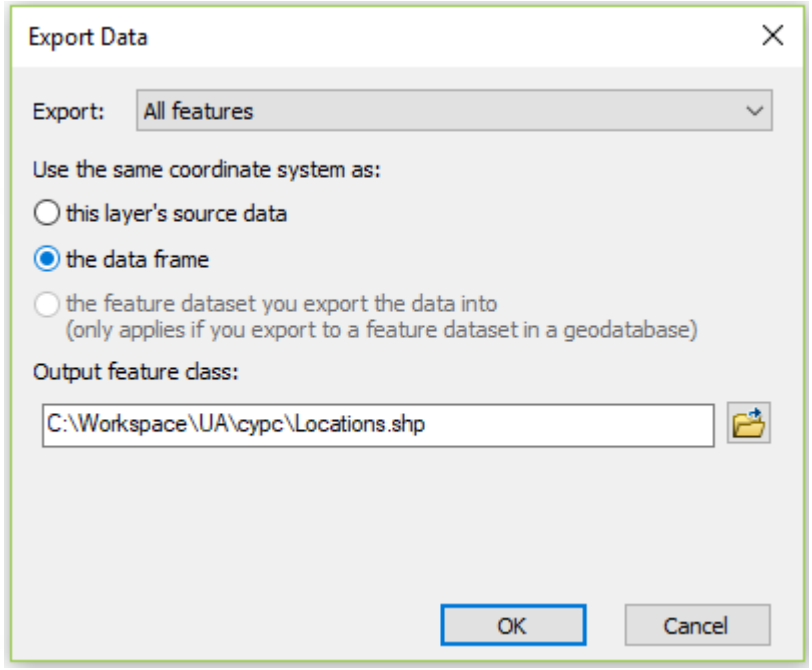


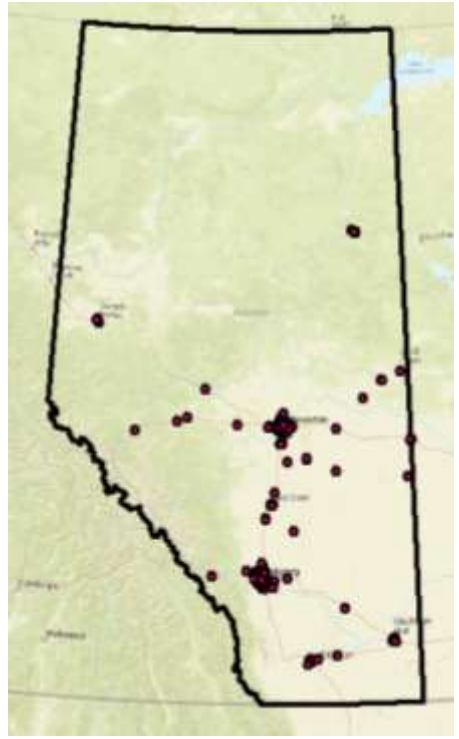


Extremely important to set the correct coordinate system

Display XY Events







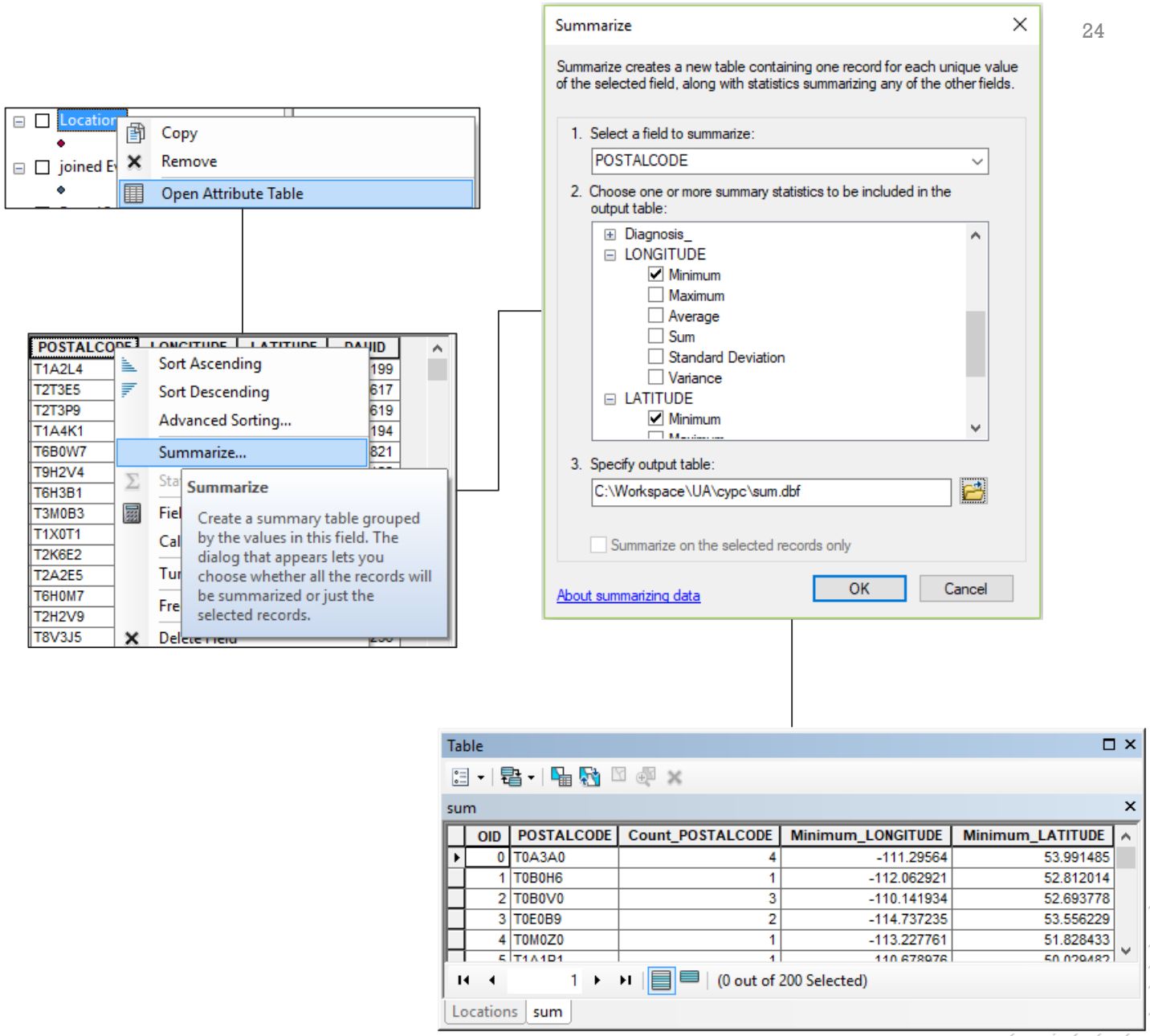
Locations.shp is a true spatial data file

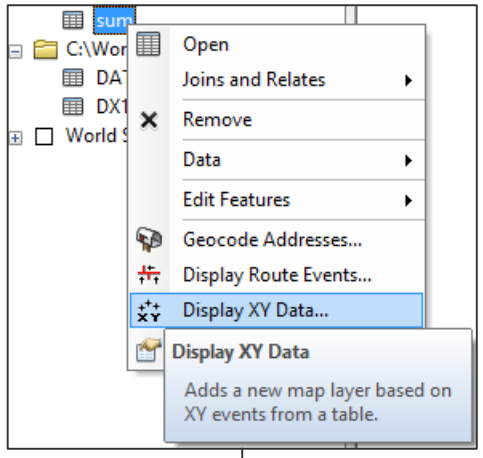
Export data

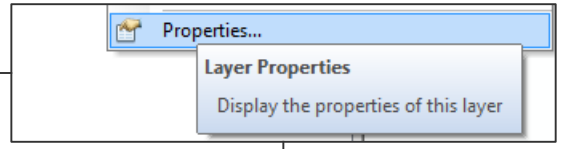
Sample analyses

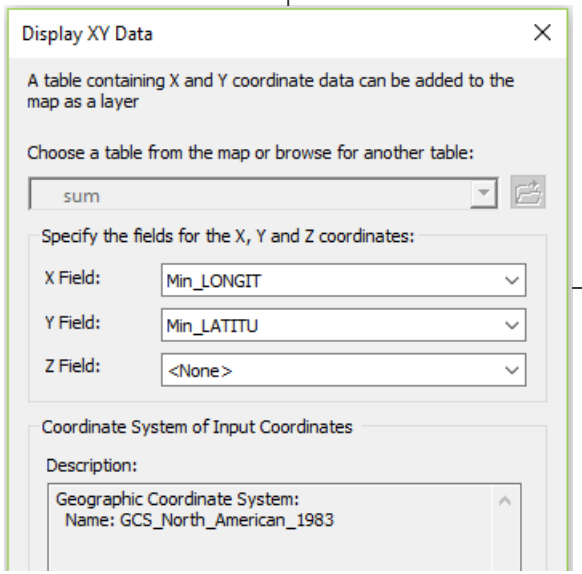
- Summary Statistics
- Overlay
- Proximity

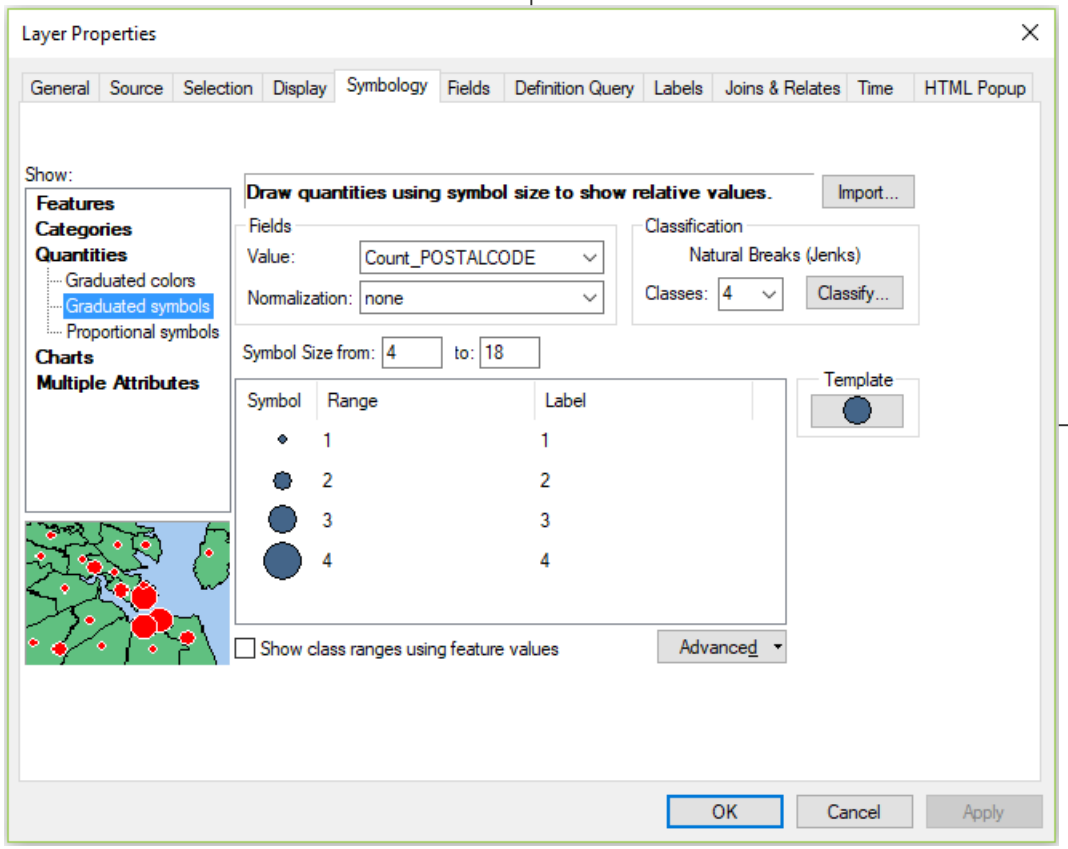
Summarize (counts)

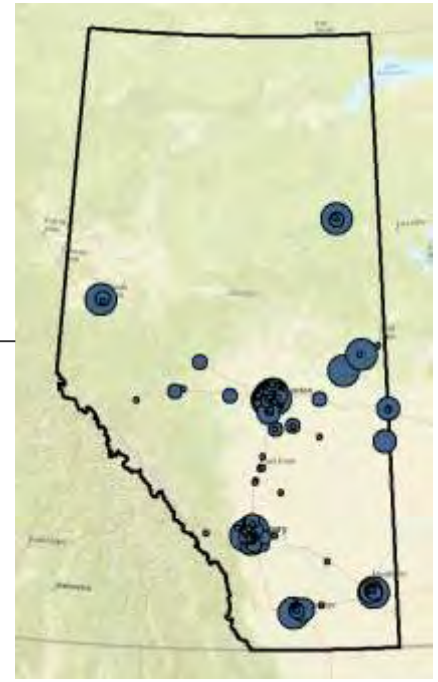




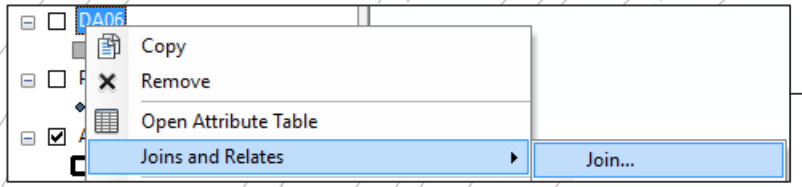




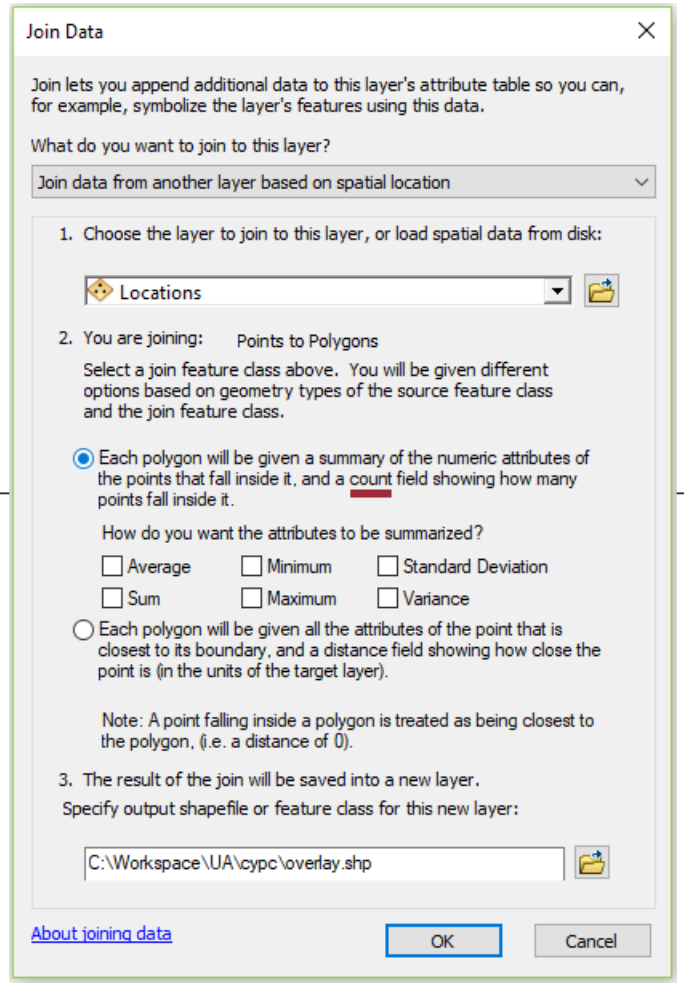


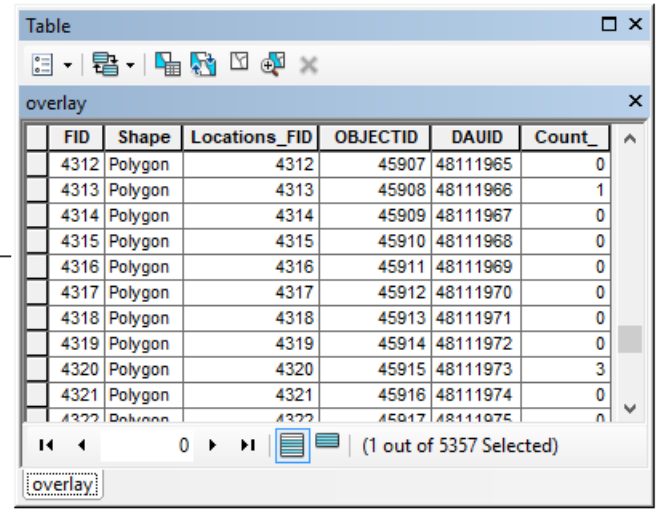


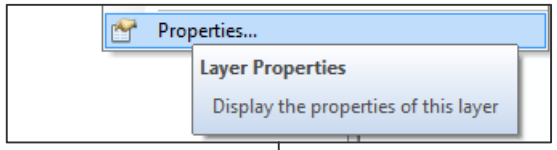
Visualize point counts

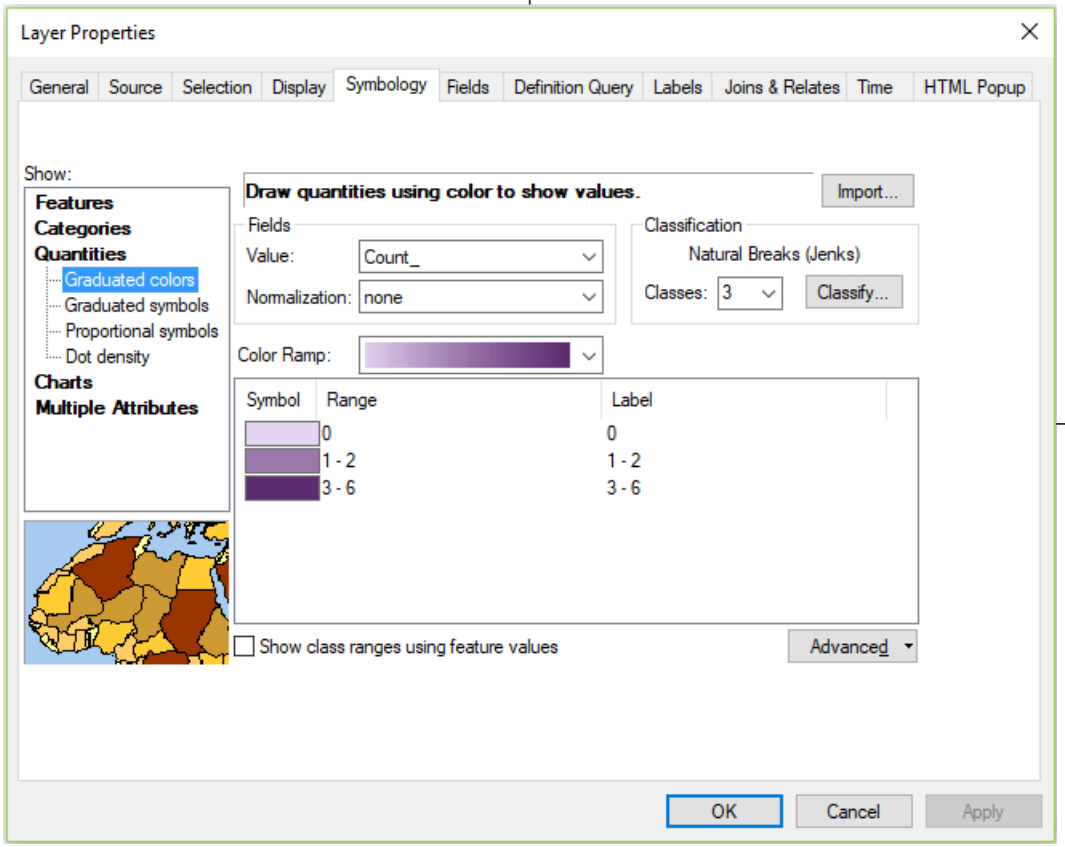


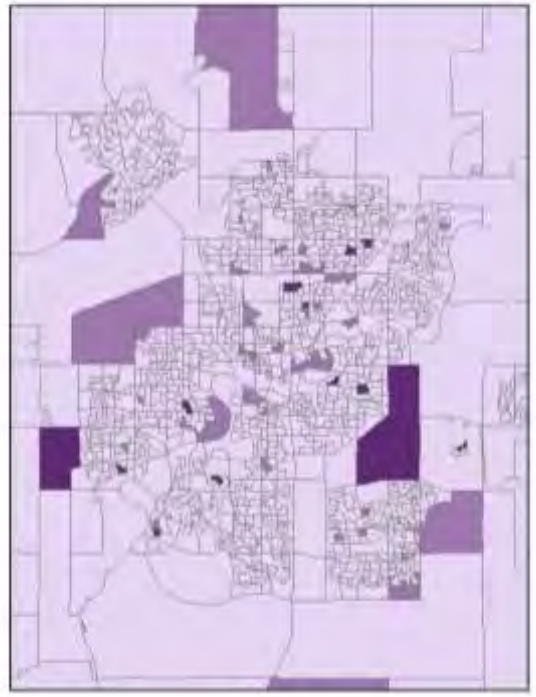
Overlay (point in polygon)





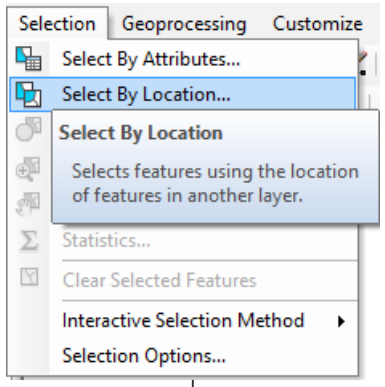


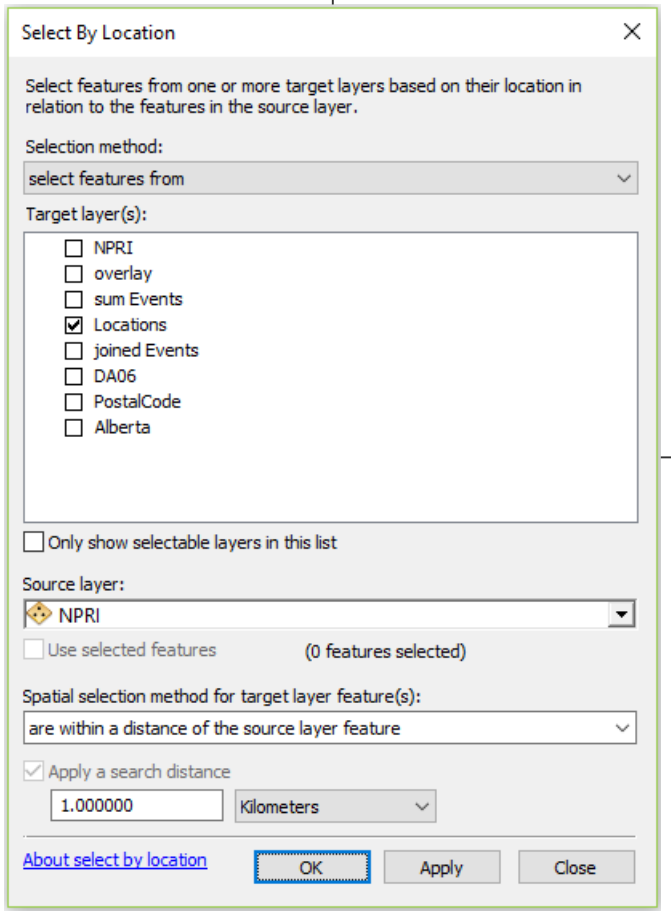


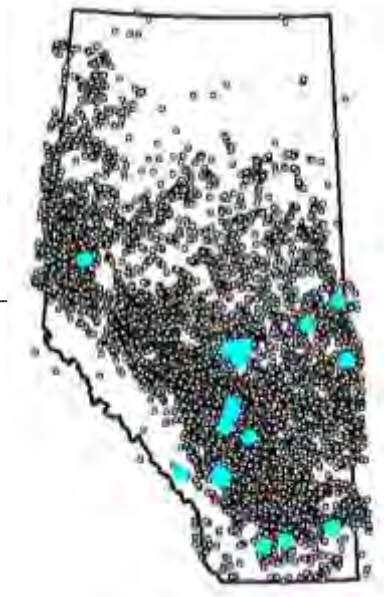


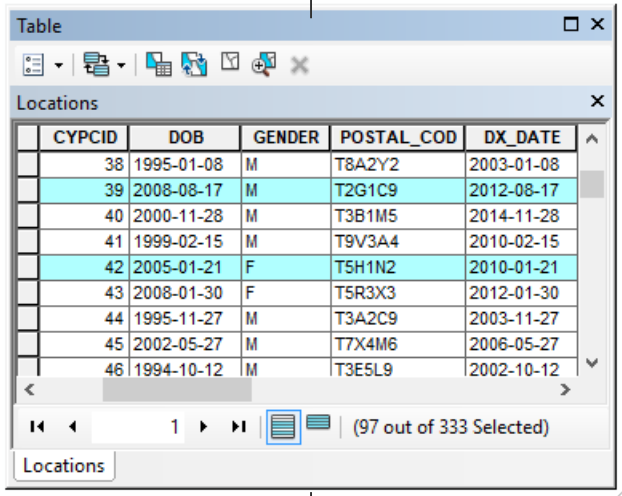
Visualize area counts

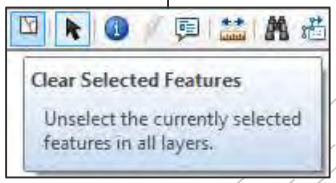
Proximity (closest features)



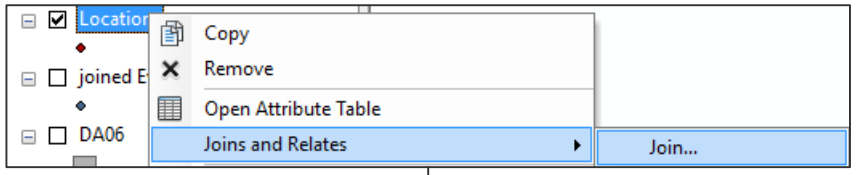


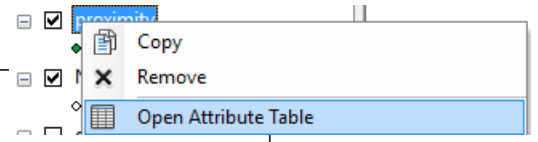


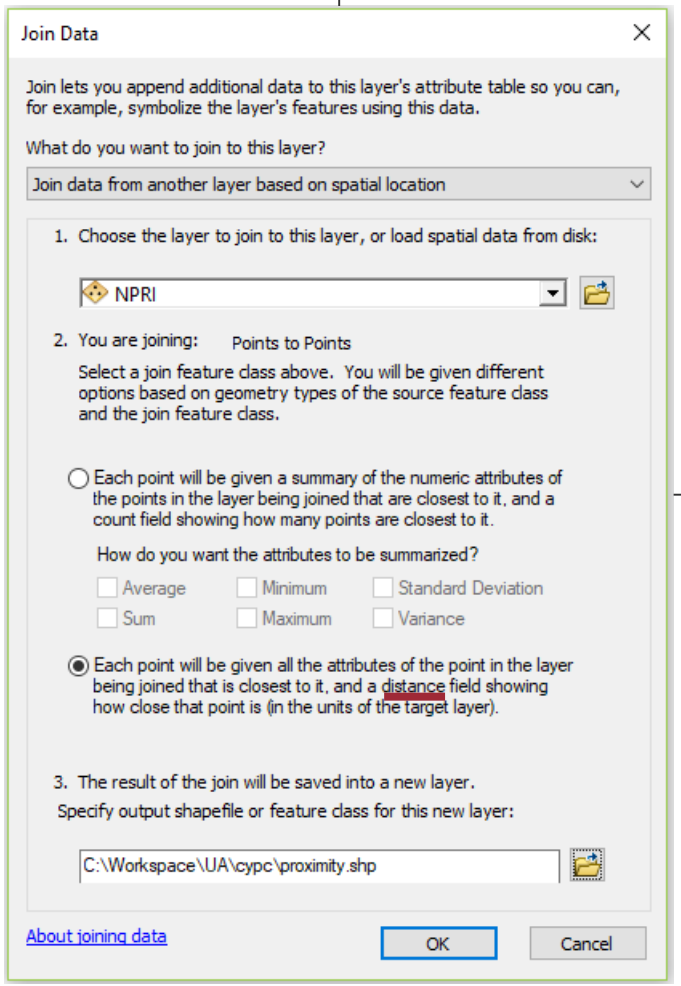


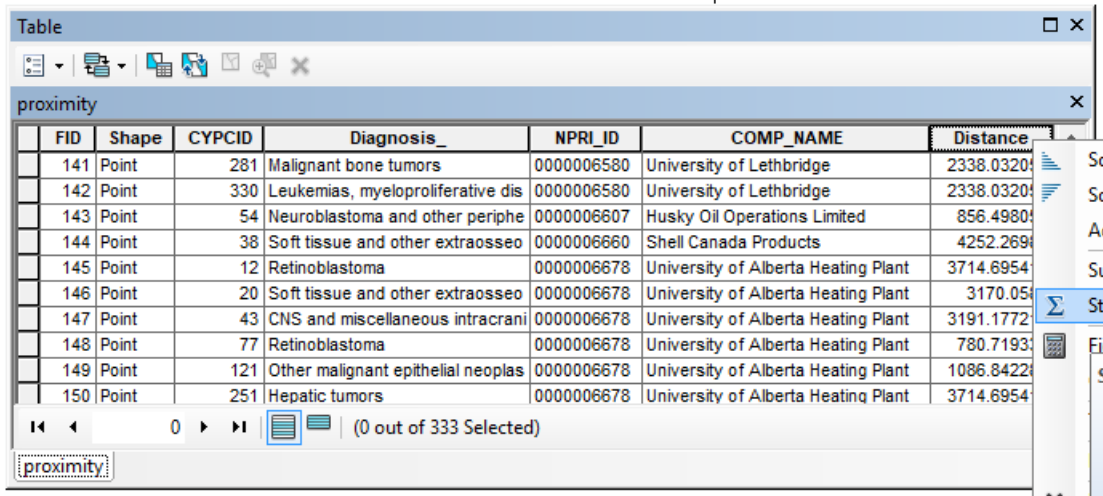


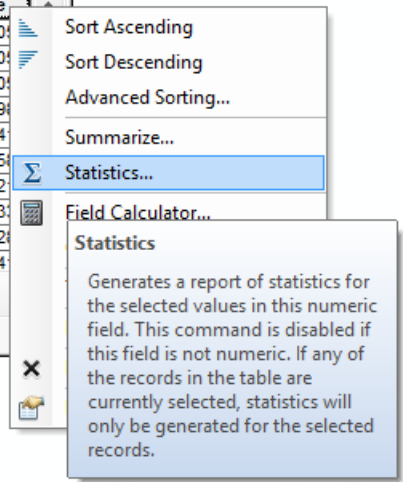
UNIVERSITY OF ALBERTA
FACULTY OF MEDICINE & DENTISTRY
Department of Pediatrics

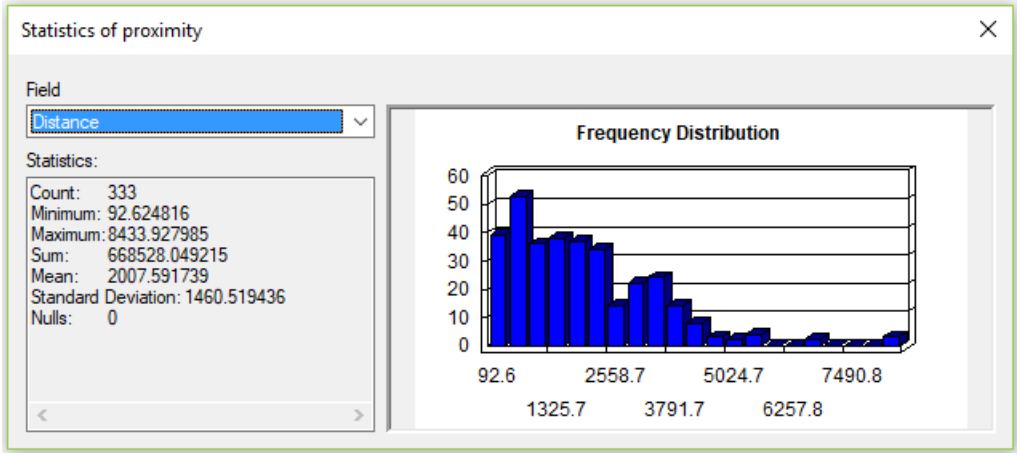












Visualize distances

GIS Demo

UNIVERSITY OF ALBERTA
FACULTY OF MEDICINE & DENTISTRY
Department of Pediatrics

More learning…

- GIS and Public Health (Cromley and McLafferty)
  https://www.library.ualberta.ca/catalog/5440524
- GIS Tutorial for Health (Kurland and Gorr)
  https://www.library.ualberta.ca/catalog/7769694
- CDC Training in GIS and Public Health
  https://www.cdc.gov/gis/gis-training.htm
- Unlimited Esri Training for the ArcGIS platform
  https://www.esri.com/training/unlimited-esri-training/

- …and a resource – SAS Bridge for Esri
  https://www.sas.com/en_ca/software/bridge-for-esri.html

Recap

- Overview of GIS
  - A system of 5 important components
- Basemaps and map projections
  - Get started with context and coordinate systems
- Geolocation
  - Locating our tabular data by postal codes
- Analyses: statistics, overlay, and proximity
  - Where we have higher/lower counts and distances to other geographic features
- Other learning references

Thank you!

Acknowledgements:

- Dr. Alvaro Osornio-Vargas
- Jesus Serrano-Lomelin
- The DoMiNO Project Team
- Randy Barber and the CYP-C Program

Cancer Research Society

READ FOR THE CURE

DoMiNO
Data Mining & Neonatal Outcomes

UNIVERSITY OF ALBERTA
FACULTY OF MEDICINE & DENTISTRY
Department of Pediatrics

UNIVERSITY OF ALBERTA
FACULTY OF SCIENCE
Department of Computing Science

UNIVERSITY OF ALBERTA
FACULTY OF SCIENCE
Department of Earth & Atmospheric Sciences